📉 Lasso Regression – Alpha Value Impact Analysis

This document summarizes the behavior of Lasso Regression with different alpha (regularization strength) values   
on the California Housing dataset. Lasso uses L1 regularization, which not only shrinks coefficients but can eliminate   
irrelevant features entirely, simplifying the model.

# 🔍 What Does Alpha Do in Lasso?

- Alpha is the regularization strength in Lasso Regression.  
- It controls how much penalty is applied to the model for using larger coefficients.  
- Increasing alpha makes the model simpler but can reduce accuracy.  
- Decreasing alpha makes the model more flexible but can lead to overfitting.  
  
Effectively:  
- Low alpha → Less regularization → More features kept  
- High alpha → More regularization → More features eliminated

# 📊 Model Comparison: Different Alpha Values

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Alpha | Features Retained | R² Score | MAE | RMSE | Interpretation |
| 0 | All (8) | 0.5758 | 0.5332 | 0.7456 | Baseline Linear Regression – Best performance |
| 0.1 | 3 | 0.4814 | 0.6221 | 0.8243 | Balanced model – few features, slight performance drop |
| 0.7 | 1 | 0.1095 | 0.8543 | 1.0802 | Too much regularization – underfitting starts |
| 1.0 | 0 | -0.0002 | 0.9061 | 1.1449 | Extreme underfitting – all features eliminated |

# 📌 Observations on Feature Selection

- With alpha = 0 (no regularization), all features are used.  
- Alpha = 0.1 keeps only 3 features: MedInc, HouseAge, and Latitude.  
- Alpha = 0.7 retains only MedInc.  
- Alpha = 1.0 eliminates all features, resulting in a useless model.  
  
This demonstrates how Lasso automatically performs feature selection.

# ✅ Conclusion

Lasso is a powerful tool for reducing model complexity and selecting important features.  
However, careful tuning of alpha is essential. Too little regularization leaves the model prone to overfitting,  
while too much causes underfitting and information loss. In this case, alpha = 0.1 strikes a practical balance between   
simplicity and performance.